Bioinformatics Analysis Combined with Experiments to Explore Neuroblastoma Prognostic Indicators Based on Immune-related Genes.

FengLing Shao

Chongqing Medical University Pediatric College: Chongqing Medical University Affiliated Children's Hospital

Zhenni Wang

Chongqing Medical University Pediatric College: Chongqing Medical University Affiliated Children's Hospital

Shan Wang (✉ wangshan@hospital.cqmu.edu.cn)

Chongqing Medical University Affiliated Children's Hospital   https://orcid.org/0000-0002-7476-1799

Bioinformatics analysis combined with experiments to explore neuroblastoma prognostic indicators based on immune-related genes.

FengLing Shao,[1] Zhenni Wang,[1] and Shan Wang[1, *]

[1]Department of Pediatric surgical oncology, Ministry of Education Key Laboratory of Child Development and Disorders, National Clinical Research Center for Child Health and Disorders, China International Science and Technology Cooperation base of Child development and Critical Disorders, Chongqing Key Laboratory of Pediatrics, Children's Hospital of Chongqing Medical University, Chongqing, PR China.
[*]This is the corresponding author.

Abstract

Background：Due to the extremely high mortality rate of children with high-risk Neuroblastoma (NB), there is an urgent need for new indicators to further classify children in the high-risk group for more precise treatment. The purpose of our research is to explore the immune-related genes in NB in the high-risk group, and to further identify and develop a prognostic nomogram based on immune IRG signatures.

Methods：Through bioinformatics analysis to explore the abnormal expression of immune-related genes in the high-risk group. Cox regression and the least absolute shrinkage and selection operator (LASSO) analysis were conducted to identify the immune and overall survival (OS) related mRNA. The accuracy of the risk score is evaluated by Kaplan-Meier method and receiver operating characteristics (ROC) analysis, which is used to build a nomogram in combination with other clinical characteristics.. Quantitative real-time polymerase chain reaction (qRT-PCR) was conducted to detect the accuracy of our results.

Results：A total of 127 common differentially expressed immune genes were found between the high-risk group and the non-high-risk group of the two data sets. Four immune-related genes (IRG) related to prognosis were identified and a risk score was established. Kaplan–Meier survival analysis and time-dependent ROC analysis showed that the 4-IRG risk score has satisfactory predictive potential and achieved consistency in the verification of external data sets. Subsequently, the risk score combined with clinical characteristics draws a nomogram. The reliability of the results was verified on 29 cases of NB tissues by qRT-PCR.

Conclusions：Overall, we have developed a powerful multi-gene classifier that can effectively classify NB patients into low- and high-risk groups with poor prognosis, and draw a nomogram for children in the high-risk group. This feature can help select high-risk patients who need more aggressive adjuvant target therapy or immunotherapy.

Keywords: bioinformatics analysis, qRT-PCR, prognostic model, immunology, high-risk neuroblastoma,

Background

Neuroblastoma (NB) is the most common extracranial solid tumor in children, accounting for 7% of all pediatric tumors in patients under 15 years of age, and 15% of all pediatric deaths caused by cancer. The world mortality rate is 0.85–1.1 per 100,000 children under 15 years of age[1]. At present, the overall survival rate (OS) of low- and medium-risk NB is above 90%[2, 3]. Despite the use of multiple treatments, the long-term survival rate of high-risk NB is less than 50%[4-6]. This may be due to the heterogeneity of NB [7-9].

The immune system plays an important role in tumorigenesis, tumor development and metastasis[10].In the past few years, the clinical application of immunotherapy in oncology has received considerable attention. So far, Anti-GD2 immunotherapy can significantly improve the event free survival (EFS) and overall survival (OS) of high-risk NB [11, 12]. CAR T Cell Therapy has also made some progress in the experimental phase [13]. Most previous studies have focused on the function of proteins in this process, but there are still relatively few further studies on the specific functions of RNA.

So far, Chen et al. have initially identified the differentially expressed genes between neuroblastoma tissue and control cell lines [14]. Wang et al. studied the prognostic significance of MYCN related genes in NB [15]. Zhong et al. did a prognostic analysis on the differentially expressed genes in stage 4 NB tissue[16]. It has not been found that there are immune-related mRNAs in the prognosis of high-risk NB. Therefore, it is important to confirm the completeness of the mRNA involved in regulating the immune response. In the immune system, mRNA is of great significance not only in providing reasonable immunotherapy methods, but also in providing accurate treatment options for tumors.

In the current study, we have explored immune-related genes from TARGET and GEO databases. In addition, we found that 4-IRG can predict survival and be used as a risk score. Kaplan–Meier survival analysis and time-dependent ROC analysis showed satisfactory prediction potential of the risk score, which was verified in the external data set. Next, the comparison of risk score and clinical factors confirmed the superiority of risk score. We combined the risk score with clinical characteristics to draw a nomogram for predicting the prognosis of high-risk children. Finally, the reliability of the results was verified on 29 cases of neuroblastoma tissues by qRT-PCR.

Methods

Gene expression and clinical data acquisition

We obtained 209 neuroblastoma gene expression data (excluding 40 samples without clear pathological diagnosis) from TARGET (https://ocg.cancer.gov/) that contained clinical information for the establishment of the model. The gene expression matrix files and clinical information of the GSE49710 and GSE73517 datasets obtained from the Gene Expression Comprehensive Library (GEO) were used for the screening of differential genes, which contained 498 and 50 samples, respectively. A dataset of 149 cases containing clinical information for external verification was downloaded from UCSC xena (http://xena.ucsc.edu/).

Identification of differential expressed immune-related genes

Use the Limma software package in R to identify the differentially expressed genes (DEG) between the high-risk group sample and the non-high-risk group. The condition for deletion is (logFC> 1, false discovery rate (FDR) <0.05). Get the list of immune genes (IRG) from the ImmPort database (https://www.immport.org). The overlapping part of DEG and IRG was selected as a collection of differentially expressed immune-related genes (DE-IRG) for further analysis and displayed by Venn diagram.

Functional analysis with DE-IRGs

Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) enrichment pathway analyses were performed to investigate the molecular functions, cellular component and biological processes of DE-IRGs. Signaling pathways that significantly related (P<0.05) to DE-IRGs were identified by DAVID (https://david.ncifcrf.gov/). Protein–protein interactions (PPI) network of the DE-IRGs were explore by The STRING database (https://string-db.org) where confidence score≥0.4 was used[17], and visualized with Cytoscape v. 3.8.12(https://cytoscape.org/). Hub nodes were identified with the Cytoscape plugin cytoHubba by the maximal clique centrality method. DE-IRGs clusters that strongly correlated in the PPI network were identified with the Cytoscape plugin MCODE. GO enrichment analysis were further performed on DE-IRGs clusters.

Identification and validation of the prognostic IRGs score

Univariate Cox analysis was used to screen DE-IRG (P<0.05) that was significantly related to overall survival (OS) in the TARGET neuroblastoma dataset. Univariate analysis results showed that meaningful genes undergo least absolute shrinkage and selection operator (LASSO) logistic regression. Subsequently, genes with non-zero coefficients were selected for multivariate Cox analysis to identify independent prognostic genes. Use these independent predictive genes (P<0.05) to further calculate the IRGs score of OS: IRG score=$\beta1* X1 + \beta1* X1 +\ldots+\beta n* Xn$ (β: coefficient derived from multiple regression; X: gene expression value). The median risk score was selected as the cutoff value of the TARGET neuroblastoma dataset, which was also used for validation in the UCSC dataset. Using Kaplan-Meier analysis, the area under the curve (AUC) of the receptor operating characteristic (ROC) curve was used to evaluate the performance of the prognostic gene signature.

Gene set enrichment and pathway analysis (GSEA)

To illustrate the biological functions of the prognostic genes in high-risk and low-risk patient groups, GSEA was performed in java GSEA (version 3.0) based on the Molecular Signatures Database version 6.2[18]. With the 209 NB samples in TARGET dataset, KEGG pathways, biological processes, cellular components, molecular functions associated with high-risk and low-risk groups were identified by using C2 (curated gene sets), C5 (GO gene sets). FDR q value < 0.05, |NES|> 1 were considered statistically significant.

Development and validation of the nomogram

The predictive factors for making a nomogram include age, stage, gender, MYCN status and IRGs scores. Concordance index and calibration curves are used to evaluate the discrimination and accuracy of the nomogram.

Validation of DKK1、HTR1A、PLXAN1 and ULBP2 expression

Use univariate cox regression to analyze DKK1, HTR1A, PLXAN1 and ULBP2 and plot their optimal cut-off value, median value, Kaplan–Meier survival analysis and ROC curve grouped by quartile values. Cbioportal (http://www.cbioportal.org/) is a web server that can analyze the relationship between DKK1, HTR1A, PLXAN1, ULBP2 and COG risk. Collect NB samples and perform quantitative real-time polymerase chain reaction (qRT-PCR) verification. After obtaining approval from the Ethics Committee of Children's Hospital of Chongqing Medical University, we collected tumor tissues from 20 high-risk children and 9 non-high-risk children. All patients were diagnosed with primary neuroblastoma and did not receive any treatment before surgery. All tumor tissues were quickly frozen in liquid nitrogen after being isolated and stored at −80°C. Total RNA was extracted from the sample using TRIzol reagent (Servicebio，Wuhan，China), and then reverse transcription and PCR reactions were performed using ReverTra Ace qPCR RT-PCR kit (Servicebio，Wuhan，China). Total RNA was extracted from the sample using TRIzol reagent (Servicebio，Wuhan，China), and then reverse transcription and PCR reactions were performed using Servicebio®RT First Strand cDNA Synthesis kit(Servicebio，Wuhan，China) and 2×SYBR Green qPCR Master Mix（Servicebio，Wuhan，China）. The sequences of the applied primers are listed in Supplementary Table S1. Then all steps were performed according to the manufacturer's instructions. Finally, the result was standardized to GAPDH and calculated by the $2^{-\Delta\Delta Ct}$ method.

Statistical analysis

The differentially expressed genes between high-risk and not high-risk tissues were analyzed by R software package "limma". Cox correlation analysis was used to confirm the relationships between the selected 4-lncRNA and patient outcomes. Survival curves were drawn with the survival package for R. $P < 0.05$ was regarded to indicate statistically significant differences.

Results

Screening of differentially expressed IRGs on neuroblastoma

The entire work flow of the research is shown in Figure 1. By comparing the expression profiles of the neuroblastoma high-risk group and non-high-risk group in the GSE49710 and GSE73517 data sets, 1569 and 1382 genes were determined to be DEG through volcano graph analysis (logFC≥1, FDR≤0.05). 2A, B). A total of 1812 immune-related genes (IRG) were downloaded from the Immport database. The 127 candidate genes defined as DE-IRGs overlapped between DEGs and IRGs and were visualized by a Venn diagram (Figure 2C). Finally, through the TARGET data set and UCSC data set, 118 DE-IRGs were determined to be used for prognostic identification.

Functional enrichment and PPI network analysis of DE-IRGs.

The functions of 127 DE-IRGs were found by analysis of KEGG and GO enrichment pathways (Figure 2D, E). Through KEGG analysis, DE-IRGs significantly enriched the biological processes related to Cytokine-cytokine receptor interaction and Hematopoietic cell lineage. GO analysis showed that DE-IRGs enriched plasma membrane, immune response and protein binding. A PPI network of the 127 DE-IRGs was established, where 125 nodes and 699 interactions were constructed, to identify the interactions between genes (Figure 3a). The top 30 candidate genes were identified to be significantly involved in the network (Figure 3B). Module analysis recognized related clustering modules in the PPI network (Figure 3D). With the DE-IRGs clusters,

GO analysis was applied for functional enrichment. The results from the PPI network and pathway analysis suggested the plasma membrane, immune response, and Chemokine signaling pathway, were densely connected and enriched in the DE-IRGs(Figure 3C,E).

Identification of DKK1、HTR1A、PLXNC1 and ULBP2 as the independent prognostic DE-IRGs

The TARGET neuroblastoma dataset (training) and the UCSC dataset (validation) were used to identify survival-related genes after 118 candidate DE-IRGs were identified. Supplementary Table 2 summarizes the clinical characteristics of these two datasets. Univariate Cox analysis was used to analyze 118 DE-IRGs in the TARGET neuroblastoma data set. Among them, 33 DE-IRGs were significantly correlated with patient survival (P <0.05) (Supplementary Table 3). Next, through LASSO penalty regression to avoid the collinearity of multiple variables, 11 (λ min=0.02614) and 5 (λ 1se=0.07273) DE-IRG survival models were obtained (Figure 4A). The coefficient of each gene in the TARGET neuroblastoma data set is shown in the figure (Figure 4B). ROC curve and box plot are used to evaluate the pros and cons of the two models (Figure 4C, D). Eleven DE-IRGs were selected, and multivariate Cox regression was performed to find out the relationship between gene expression and patient OS. Among them, DKK1 (HR = 0.64, 95% CI 0.44-0.92, P = 0.01616), HTR1A (HR = 2.54) , 95% CI 1.31-4.93, P = 0.006), PLXNC1 (HR = 0.66, 95% CI 0.48-0.90, P = 0.00918) and ULBP2 (HR = 0.57, 95% CI 0.33-0.96, P = 0.03313) are independent prognostic genes (Figure 4E).

Development and validation of IRGs score model

The independent prognostic genes DKK1、HTR1A、PLXNC1 and ULBP2 were chosen to establish a risk score model. From multivariate Cox regression, the coefficients for DKK1、HTR1A、PLXNC1 and ULBP2 were -0.44768、0.93075、-0.42222 and -0.57078 respectively. Therefore, the IRG score of each patient was calculated according to the formula: IRGs score = (-0.44768) × (expression value of DKK1) + 0.93075 × (expression value of HTR1A)+ (-0.42222) × (expression value of PLXNC1) + (-0.57078) × (expression value of ULBP2) . According to the median risk score (-3.11074) in the TARGET neuroblastoma dataset, patients were divided into high-risk and low-risk groups. Kaplan-Meier survival analysis showed that the OS of patients in the high-risk group was significantly shorter than that of patients in the low-risk group (Figure 5A). The ROC curve was used to evaluate the 1-, 3-, and 5-year prediction performance of the risk score (Figure 5C). The results showed that the 1-, 3-, and 5-year AUC values were 0.738, 0.719, and 0.688, respectively. The distribution of risk score, OS, DKK1, HTR1A, PLXNC1, and ULBP2 expression is also shown, which indicates that patients with high risk scores have lower expressions of DKK1, PLXNC1, and ULBP2, higher HTR1A expression and more deaths (Figure 5E). In addition, in order to confirm that the risk score formula is better than other commonly used clinical parameters (gender, age at diagnosis, INSS staging, MYCN amplification status), we used the roc curve to compare their 135-year overall survival rate and the predictive performance of event-free survival rate. Results It shows that under all conditions, the risk score is better than other parameters (Supplementary Figure 1A, B).

According to the same formula and the above cut-off value, patients in the UCSC data set used for external verification were divided into high-risk and low-risk groups for comparison. Consistent with the previous results, the survival rate of high-risk patients was significantly lower than that of low-risk patients (Figure 5B). The 1-, 3-, and 5-year AUC values of the ROC working

curve were 0.606, 0.676, and 0.620, respectively (Figure 5D). In addition, the expression of DKK1, PLXNC1 and ULBP2 in the high-risk group was reduced, and the number of deaths in the high-risk group was also higher than that in the low-risk group in the UCSC validation data set (Figure 5F).

Gene set enrichment and pathway analysis for DE-IRGs

To investigate the underlying molecular mechanism of the IRGs signature, we conducted GSEA comparing the high-risk group with the low-risk group in 209 NB patients from TARGET. The high-risk group has only 1 rich GO pathway and 1 rich KEGG pathway. However, 148 and 25 pathways in the GO and KEGG analysis were determined to be related to the low-risk group, and the top 10 important terms for each module are summarized in Supplementary Table S4. The results showed that the high-risk group may be related to KERATIN FILAMENT and MATURITY ONSET DIABETES OF THE YOUNG, and the low-risk group may be related to vacuum membrane, go ligand dependent nuclear receptor transcription coactivator activity and neurotrophin signaling pathway (Supplementary Figure 2).

Development and evaluation of nomogram based on IRG and clinicopathological risk factors

In order to make prognostic judgments for high-risk neuroblastoma, we deleted 30 non-high-risk samples from the TARGET data set and included 179 COG high-risk group samples to construct a nomogram. Prognostic factors include gender, age at diagnosis, INSS stage, MYCN amplification status, and risk score to construct a nomogram to predict 1-year, 3-year, 5-year overall survival rate and median overall survival time (Figure 6A, B) . The concordance index of the nomogram is 0.633. The calibration curve shows that the nomogram has good agreement between the predictions of 1-year, 3-year, and 5-year overall survival rates and actual observations (Figure 6C). However, when predicting the three-year overall survival rate, the risk of the low-risk group will be underestimated.

QRT-PCR verification of DKK1, HTR1A, PLXNC1 and ULBP2 and evaluation of independent predictive ability

Univariate analysis of DKK1, HTR1A, PLXNC1, and ULBP2 showed that although these four genes performed poorly in the Kaplan–Meier survival analysis grouped by the median, they were at the best cutoff value and quartile Kaplan–Meier survival analysis showed a good prognostic ability (Figure 7A). We use qRT-PCR to further verify their expression in clinical tissues. Interestingly, compared with the tumor tissues of the low-risk group, the relative expression levels of signature genes in the tumor tissues of the high-risk group were significantly reduced (Figure 7B). In addition, the relationship between the expression levels of DKK1, HTR1A, PLXNC1 and ULBP2 and the risk of COG in Cbioportal is also roughly in line with the results of qRT-PCR (Figure 7C).

Discussion

Neuroblastoma (NB) is the cancer with the highest incidence of solid tumors in children, and the overall survival rate of high-risk children is less than 50%[19]. It is urgent to identify prognostic biomarkers to personalize the treatment plan for high-risk children. With the continuous exploration of the role of the immune system in tumorigenesis, the phenomenon that the immune system is activated in tumors and then further mediates immune cells to kill tumor cells has been widely recognized [20]. The expression of IRGs in tumors directly reflects the

status of the tumor immune system. Therefore, the use of IRG as a prognostic factor for tumors is convincing. Our study is the first to establish a prognostic model for NB patients based on IRGs. With the development and wide application of next-generation" sequencing technology, NB has also established many other genetic prognosis models, such as those based on MYCN-related genes or INSS 4 stage-related genes[15, 16]. Unlike other studies, we mainly focus on changes in the expression of IRGs and construct prognostic models for high-risk patients. In addition to having excellent prognostic value for the entire NB population, our research can also guide the clinical exploration of individualized treatment and immunotherapy for children with NB in the high-risk group to a certain extent.

We used public databases and bioinformatics methods to screen the most influential immune-related genes that affect patient survival. A total of nearly 1,000 samples (n = 962) were recruited. As far as we know, this is the largest cohort to establish a prognostic scoring system based on immune-related gene signatures in NB. In order to ensure the consistency between different data sets, we only use the gene microarray data generated by the GPL16876 platform (Agilent-020382 Human Custom Microarray 44k), so the heterogeneity within the cohort is eliminated to some extent. Based on machine learning algorithms, we have established a risk scoring system based on 4 genes to individually predict the prognosis of NB patients. The expression of these 4 genes has been verified on clinical organizations and external data sets. Combining several clinical parameters, we generated a nomogram that can predict the prognosis of high-risk patients. It is worth noting that the expression of the four genes has been verified in our clinical tissues, which means that the four genes in our model have better credibility. On the other hand, the prognostic factors in our nomogram are easy to obtain in clinical work, so these accessible parameters can be used to achieve relatively accurate prognostic survival predictions.

In this study, the expression of DKK1 was negatively correlated with poor prognosis. As a typical inhibitor of wnt/β-catenin, DKK1 plays a key role in tumor pathogenesis[21]. In a variety of cancers, DKK1 regulates tumor angiogenesis, mediates cancer metastasis, regulates chemoresistance, and shows potential as a serum biomarker for diagnosis and prognosis[22-25]. According to reports, in neuroblastoma, DKK1 is down-regulated by MYCN, and its expansion and overexpression are poor prognostic factors for neuroblastoma [26]. However, DKK1 is down-regulated by XIST through EZH2 induced H3 histone methylation, thereby promoting the growth, migration and invasion of neuroblastoma cells, and delaying tumor progression. PMID: 31208278. In addition, DKK1 is released by neuroblastoma cells, which can affect the balance between osteoblast production and osteoclast production, thereby contributing to the occurrence of osteolytic metastasis [27]. Our results show that higher expression of DKK1 indicates a higher survival time for patients, which has been confirmed by external verification. This indicates that DKK1 in NB may play a protective role as a whole, and the mechanism of DKK1 in NB deserves further study.

The HTR1A gene encodes the 5-HT1A receptor protein, which is one of the seven inhibitory G protein-coupled serotonin receptors[28]. In the past few years, studies have found that 5-HT1A receptor protein is involved in neurodevelopment [29] and is related to neuronal apoptosis [30]. This receptor protein plays a key role in neuron migration, axon growth and synapse formation, and affects the process of neurodevelopment [31]. On the one hand, 5-HT1A has a growth-promoting effect on carcinoid and aggressive cancers. It can induce tumor growth and survival of several cell types, including pancreas, liver, breast, bladder, lung, prostate and other

cells[32]. On the other hand, the high expression of 5-HT1A is not only detected in prostate cancer cell lines and small cell lung cancer cell lines, but also in the pathological tissues of prostate cancer, bladder cancer and breast cancer patients by immunohistochemical staining. A positive result of 5-HT1A is given [33-36]. For now, there are few studies on HTR1A in NB. For example, Murata et al. found DNA methylation of HTR1A in neuroblastoma cell lines [37]. Bence et al. found that HTR1A may be related to transcriptional modulation of monoaminergic neurotransmission genes [38]. The above results suggest that HTR1A is related to the poor prognosis of tumors, which is also consistent with the result that HTR1A is a poor prognostic factor in NB, and HTR1A may be a potential prognostic biomolecular marker for NB.

PLXNC1 is a member of the transmembrane receptor family, and the corresponding gene is located on chromosome 12 in the human genome. Plexins are a part of the semaphorin gene superfamily, which includes the semaphorins and the receptor tyrosine kinases MET and RON[39]. Granja and König suggested that PLXNC1 plays an important role in inflammation, especially acute inflammation [40, 41]. In the research on tumors, on the one hand, PLXNC1 was identified as a protective gene in acute myeloid leukemia, papillary thyroid cancer, and melanoma[42-44]. On the other hand, PLXNC1 is considered to be related to the poor prognosis of gastric cancer and liver cancer[45, 46], and was mainly related to the aggressive phenotype in a pan-cancer study[47]. There are few studies on PLXNC1 in NB. Lebedev reported that PLXNC1 gene expression is correlated with the expression levels of TrkA or KIT in AML and NB, and has prognostic value for both cancers[48]. In our study, PLXNC1 was identified as a protective gene, which provides a reference for the clinical evaluation of patients. On the other hand, there are also research results showing that PLXNC1 is related to the poor prognosis of some cancers, which may promote the study of its adverse effects in neuroblastoma. In summary, PLXNC1 can indicate patient survival and is a potential biomarker for neuroblastoma.

ULBP2 is a NKG2D ligand, highly expressed in transformed cells [49] and is believed to provide a ligand for NK cells, which is conducive to tumor cell death[50-52]. However, studies have shown that overexpression of ULBP2 is associated with poor prognosis of cancers, including expression in melanoma, non-small cell lung cancer and ovarian cancer[53-56]. Hayakawa et al. believed that the upregulation of ULBP2 would lead to a bad result may be due to the secretion or cleavage of this protein is related to reduced NK mediated tumor lysis[52]. In the study of ULBP2 in NB, the expression of ULBP2 was detected in half of the cases of primary NB, but only in a few cell lines in NB cell lines[57]. The difference in the expression pattern between the primary tumor and the cell line may be related to the changes caused by the long-term culture of the cells or the sensitivity of the detection technology used. Our research shows that ULBP2 plays a protective role in NB, and based on current knowledge, the specific function of the interaction between ULBP2 and NK is still unclear.

To be sure, our analysis has some limitations. First of all, our analysis is based on expression at the mRNA level, without considering the expression at the protein level or post-transcriptional modification, which also has important biological effects. In addition, our IRGs are screened in the differential genes between high-risk patients and non-high-risk patients. However, we may have overlooked some genes with key prognostic value, because these genes are not necessarily different between children in high-risk groups and children in non-high-risk groups. Finally, there is a lack of validation in more neuroblastoma cohorts, which is limited by the availability of data.

Conclusions

In summary, our team established a risk scoring formula based on 4 immune-related gene markers in neuroblastoma, where a high score was significantly associated with a poor prognosis, which was verified in an external cohort. Combining several other clinical information to draw a nomogram, it can be used as a powerful and accurate tool to predict the survival rate of high-risk patients. It can help clinicians to more accurately choose the best treatment plan for NB patients.

Supplementary Material

Supplementary Table S1: Primers applied in qRT-PCR.

Supplementary Table S2: Summary of clinical character from TARGET dataset and UCSC dataset.

Supplementary Table S3: Univariate Cox regression analysis of DE-IRGs and overall survival in TARGET neuroblastoma cohort.

Supplementary Table S4: Enrichment for High-risk patients in GO analysis.

Supplementary Figure S1: ROC curve of risk score, INSS staging, MYCN status, age at diagnosis and COG risk stratification for 1-year, 3-year, 5-year overall survival and event-free survival.

Supplementary Figure S2: Gene set enrichment and pathway analysis (GSEA).

Abbreviations

NB: Neuroblastoma; GEO：Gene Expression Omnibus database; LASSO: Least absolute shrinkage and selection operator; ROC: Receiver operating characteristic; DEG: Differentially expressed gene; IRG: Immune related genes; GSEA: Gene Set Enrichment Analysis;OS: Overall survival; qRT-PCR: Quantitative real-time polymerase chain reaction.

Authors' contributions

Shao performed the research, analyzed the data, and wrote the manuscript, Wang collected the clinical samples, and Wang[*] contributed to carrying out additional analyses, discussed the results, provided scientific advice, and revised the manuscript. All authors read and approved the final manuscript.

Data availability statement

The data supporting the findings of this study are available from the corresponding author upon request.

Ethics approval and consent to participate

The entire study was approved by the Medical Ethics Committee of Children's Hospital of Chongqing Medical University.

Consent for publication

Not applicable.

Competing interests

There are no conflicts of interest to be reported.

Author Disclosure Statement

No competing financial interests exist.

References

1. Kholodenko IV, Kalinovsky DV, Doronin, II, Deyev SM, Kholodenko RV: **Neuroblastoma Origin and Therapeutic Targets for Immunotherapy**. *J Immunol Res* 2018, **2018**:7394268.

2. Strother DR, London WB, Schmidt ML, Brodeur GM, Shimada H, Thorner P, Collins MH, Tagge E, Adkins S, Reynolds CP *et al*: **Outcome after surgery alone or with restricted use of chemotherapy for patients with low-risk neuroblastoma: results of Children's Oncology Group study P9641**. *J Clin Oncol* 2012, **30**(15):1842-1848.

3. Baker DL, Schmidt ML, Cohn SL, Maris JM, London WB, Buxton A, Stram D, Castleberry RP, Shimada H, Sandler A *et al*: **Outcome after reduced chemotherapy for intermediate-risk neuroblastoma**. *N Engl J Med* 2010, **363**(14):1313-1323.

4. Pearson AD, Pinkerton CR, Lewis IJ, Imeson J, Ellershaw C, Machin D, European Neuroblastoma Study G, Children's C, Leukaemia G: **High-dose rapid and standard induction chemotherapy for patients aged over 1 year with stage 4 neuroblastoma: a randomised trial**. *Lancet Oncol* 2008, **9**(3):247-256.

5. Matthay KK, Reynolds CP, Seeger RC, Shimada H, Adkins ES, Haas-Kogan D, Gerbing RB, London WB, Villablanca JG: **Long-term results for children with high-risk neuroblastoma treated on a randomized trial of myeloablative therapy followed by 13-cis-retinoic acid: a children's oncology group study**. *J Clin Oncol* 2009, **27**(7):1007-1013.

6. Holmes K, Potschger U, Pearson ADJ, Sarnacki S, Cecchetto G, Gomez-Chacon J, Squire R, Freud E, Bysiek A, Matthyssens LE *et al*: **Influence of Surgical Excision on the Survival of Patients With Stage 4 High-Risk Neuroblastoma: A Report From the HR-NBL1/SIOPEN Study**. *J Clin Oncol* 2020, **38**(25):2902-2915.

7. Pinto NR, Applebaum MA, Volchenboum SL, Matthay KK, London WB, Ambros PF, Nakagawara A, Berthold F, Schleiermacher G, Park JR *et al*: **Advances in Risk Classification and Treatment Strategies for Neuroblastoma**. *J Clin Oncol* 2015, **33**(27):3008-3017.

8. Matthay KK, Maris JM, Schleiermacher G, Nakagawara A, Mackall CL, Diller L, Weiss WA: **Neuroblastoma**. *Nat Rev Dis Primers* 2016, **2**:16078.

9. Park JR, Bagatell R, London WB, Maris JM, Cohn SL, Mattay KK, Hogarty M, Committee COGN: **Children's Oncology Group's 2013 blueprint for research: neuroblastoma**. *Pediatr Blood Cancer* 2013, **60**(6):985-993.

10. Chen DS, Mellman I: **Elements of cancer immunity and the cancer-immune set point**. *Nature* 2017, **541**(7637):321-330.

11. Yu AL, Gilman AL, Ozkaynak MF, London WB, Kreissman SG, Chen HX, Smith M, Anderson B, Villablanca JG, Matthay KK *et al*: **Anti-GD2 antibody with GM-CSF, interleukin-2, and isotretinoin for neuroblastoma**. *N Engl J Med* 2010, **363**(14):1324-1334.

12. Cheung NK, Cheung IY, Kushner BH, Ostrovnaya I, Chamberlain E, Kramer K, Modak S: **Murine anti-GD2 monoclonal antibody 3F8 combined with granulocyte-macrophage colony-stimulating factor and 13-cis-retinoic acid in high-risk patients with stage 4 neuroblastoma in first remission**. *J Clin Oncol* 2012, **30**(26):3264-3270.

13. Richards RM, Sotillo E, Majzner RG: **CAR T Cell Therapy for Neuroblastoma**. *Front Immunol* 2018, **9**:2380.

14. Chen B, Ding P, Hua Z, Qin X, Li Z: **Analysis and identification of novel biomarkers involved in neuroblastoma via integrated bioinformatics**. *Invest New Drugs* 2020.

15. Wang H, Wang X, Xu L, Zhang J, Cao H: **Prognostic significance of MYCN related genes in pediatric neuroblastoma: a study based on TARGET and GEO datasets**. *BMC Pediatr* 2020, **20**(1):314.

16. Zhong X, Liu Y, Liu H, Zhang Y, Wang L, Zhang H: **Identification of Potential Prognostic Genes for Neuroblastoma**. *Front Genet* 2018, **9**:589.

17. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP *et al*: **STRING v10: protein-protein interaction networks, integrated over the tree of life**. *Nucleic Acids Res* 2015, **43**(Database issue):D447-452.

18. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al*: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles**. *Proc Natl Acad Sci U S A* 2005, **102**(43):15545-15550.

19. Maris JM, Hogarty MD, Bagatell R, Cohn SL: **Neuroblastoma**. *Lancet* 2007, **369**(9579):2106-2120.

20. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD *et al*: **The prognostic landscape of genes and infiltrating immune cells across human cancers**. *Nat Med* 2015, **21**(8):938-945.

21. Mariz K, Ingolf JB, Daniel H, Teresa NJ, Erich-Franz S: **The Wnt inhibitor dickkopf-1: a link between breast cancer and bone metastases**. *Clin Exp Metastasis* 2015, **32**(8):857-866.

22. Park H, Jung HY, Choi HJ, Kim DY, Yoo JY, Yun CO, Min JK, Kim YM, Kwon YG: **Distinct roles of DKK1 and DKK2 in tumor angiogenesis**. *Angiogenesis* 2014, **17**(1):221-234.

23. Hu P, Chu J, Wu Y, Sun L, Lv X, Zhu Y, Li J, Guo Q, Gong C, Liu B *et al*: **NBAT1 suppresses breast cancer metastasis by regulating DKK1 via PRC2**. *Oncotarget* 2015, **6**(32):32410-32425.

24. Jia X, Li N, Peng C, Deng Y, Wang J, Deng M, Lu M, Yin J, Zheng G, Liu H *et al*: **miR-493 mediated DKK1 down-regulation confers proliferation, invasion and chemo-resistance in gastric cancer cells**. *Oncotarget* 2016, **7**(6):7044-7054.

25. Han SX, Zhou X, Sui X, He CC, Cai MJ, Ma JL, Zhang YY, Zhou CY, Ma CX, Varela-Ramirez A *et al*: **Serum dickkopf-1 is a novel serological biomarker for the diagnosis and prognosis of pancreatic cancer**. *Oncotarget* 2015, **6**(23):19907-19917.

26. Koppen A, Ait-Aissa R, Hopman S, Koster J, Haneveld F, Versteeg R, Valentijn LJ: **Dickkopf-1 is down-regulated by MYCN and inhibits neuroblastoma cell proliferation**. *Cancer Lett* 2007, **256**(2):218-228.

27. Granchi D, Corrias MV, Garaventa A, Baglio SR, Cangemi G, Carlini B, Paolucci P, Giunti A,

Baldini N: **Neuroblastoma and bone metastases: clinical significance and prognostic value of Dickkopf 1 plasma levels**. *Bone* 2011, **48**(1):152-159.

28. Albert PR, Le Francois B, Vahid-Ansari F: **Genetic, epigenetic and posttranscriptional mechanisms for treatment of major depression: the 5-HT1A receptor gene as a paradigm**. *J Psychiatry Neurosci* 2019, **44**(3):164-176.

29. del Olmo E, Lopez-Gimenez JF, Vilaro MT, Mengod G, Palacios JM, Pazos A: **Early localization of mRNA coding for 5-HT1A receptors in human brain during development**. *Brain Res Mol Brain Res* 1998, **60**(1):123-126.

30. Singh JK, Chromy BA, Boyers MJ, Dawson G, Banerjee P: **Induction of the serotonin1A receptor in neuronal cells during prolonged stress and degeneration**. *J Neurochem* 1996, **66**(6):2361-2372.

31. Hirst WD, Andree TH, Aschmies S, Childers WE, Comery TA, Dawson LA, Day M, Feingold IB, Grauer SM, Harrison BL *et al*: **Correlating efficacy in rodent cognition models with in vivo 5-hydroxytryptamine1a receptor occupancy by a novel antagonist, (R)-N-(2-methyl-(4-indolyl-1-piperazinyl)ethyl)-N-(2-pyridinyl)-cyclohexane carboxamide (WAY-101405)**. *J Pharmacol Exp Ther* 2008, **325**(1):134-145.

32. Corvino A, Fiorino F, Severino B, Saccone I, Frecentese F, Perissutti E, Di Vaio P, Santagada V, Caliendo G, Magli E: **The Role of 5-HT1A Receptor in Cancer as a New Opportunity in Medicinal Chemistry**. *Curr Med Chem* 2018, **25**(27):3214-3227.

33. Abdul M, Logothetis CJ, Hoosein NM: **Growth-inhibitory effects of serotonin uptake inhibitors on human prostate carcinoma cell lines**. *J Urol* 1995, **154**(1):247-250.

34. Mills SE, Wolfe JT, 3rd, Weiss MA, Swanson PE, Wick MR, Fowler JE, Jr., Young RH: **Small cell undifferentiated carcinoma of the urinary bladder. A light-microscopic, immunocytochemical, and ultrastructural study of 12 cases**. *Am J Surg Pathol* 1987, **11**(8):606-617.

35. Cattaneo MG, Codignola A, Vicentini LM, Clementi F, Sher E: **Nicotine stimulates a serotonergic autocrine loop in human small-cell lung carcinoma**. *Cancer Res* 1993, **53**(22):5566-5568.

36. Kopparapu PK, Tinzl M, Anagnostaki L, Persson JL, Dizeyi N: **Expression and localization of serotonin receptors in human breast cancer**. *Anticancer Res* 2013, **33**(2):363-370.

37. Murata Y, Bundo M, Sunaga F, Kasai K, Iwamoto K: **DNA Methylation Profiling in a Neuroblastoma Cell Line Exposed to the Antipsychotic Perospirone**. *Pharmacopsychiatry* 2019, **52**(2):63-69.

38. Bence M, Koller J, Sasvari-Szekely M, Keszler G: **Transcriptional modulation of monoaminergic neurotransmission genes by the histone deacetylase inhibitor trichostatin A in neuroblastoma cells**. *J Neural Transm (Vienna)* 2012, **119**(1):17-24.

39. Tamagnone L, Comoglio PM: **Signalling by semaphorin receptors: cell guidance and beyond**. *Trends Cell Biol* 2000, **10**(9):377-383.

40. Granja T, Kohler D, Mirakaj V, Nelson E, Konig K, Rosenberger P: **Crucial role of Plexin C1 for pulmonary inflammation and survival during lung injury**. *Mucosal Immunol* 2014, **7**(4):879-891.

41. Konig K, Marth L, Roissant J, Granja T, Jennewein C, Devanathan V, Schneider M, Kohler D, Zarbock A, Rosenberger P: **The plexin C1 receptor promotes acute inflammation**. *Eur J Immunol* 2014, **44**(9):2648-2658.

42. Stirewalt DL, Meshinchi S, Kopecky KJ, Fan W, Pogosova-Agadjanyan EL, Engel JH, Cronk MR, Dorcy KS, McQuary AR, Hockenbery D *et al*: **Identification of genes with abnormal expression changes in acute myeloid leukemia**. *Genes Chromosomes Cancer* 2008, **47**(1):8-20.

43. Kumasaka MY, Yajima I, Iida M, Takahashi H, Inoue Y, Fukushima S, Ihn H, Takeda K, Naito Y, Yoshikawa T *et al*: **Correlated expression levels of endothelin receptor B and Plexin C1 in melanoma**. *Am J Cancer Res* 2015, **5**(3):1117-1123.

44. Wang Z, Wang X, Zhou H, Dan X, Jiang L, Wu Y: **Long non-coding RNA CASC2 inhibits tumorigenesis via the miR-181a/PLXNC1 axis in melanoma**. *Acta Biochim Biophys Sin (Shanghai)* 2018, **50**(3):263-272.

45. Chen J, Liu H, Chen J, Sun B, Wu J, Du C: **PLXNC1 Enhances Carcinogenesis Through Transcriptional Activation of IL6ST in Gastric Cancer**. *Front Oncol* 2020, **10**:33.

46. Odabas G, Cetin M, Turhal S, Baloglu H, Sayan AE, Yagci T: **Plexin C1 Marks Liver Cancer Cells with Epithelial Phenotype and Is Overexpressed in Hepatocellular Carcinoma**. *Can J Gastroenterol Hepatol* 2018, **2018**:4040787.

47. Zhang X, Shao S, Li L: **Characterization of Class-3 Semaphorin Receptors, Neuropilins and Plexins, as Therapeutic Targets in a Pan-Cancer Study**. *Cancers (Basel)* 2020, **12**(7).

48. Lebedev TD, Vagapova ER, Popenko VI, Leonova OG, Spirin PV, Prassolov VS: **Two Receptors, Two Isoforms, Two Cancers: Comprehensive Analysis of KIT and TrkA Expression in Neuroblastoma and Acute Myeloid Leukemia**. *Front Oncol* 2019, **9**:1046.

49. Groh V, Rhinehart R, Secrist H, Bauer S, Grabstein KH, Spies T: **Broad tumor-associated expression and recognition by tumor-derived gamma delta T cells of MICA and MICB**. *Proc Natl Acad Sci U S A* 1999, **96**(12):6879-6884.

50. Cosman D, Mullberg J, Sutherland CL, Chin W, Armitage R, Fanslow W, Kubin M, Chalupny NJ: **ULBPs, novel MHC class I-related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the NKG2D receptor**. *Immunity* 2001, **14**(2):123-133.

51. Ohashi M, Eagle RA, Trowsdale J: **Post-translational modification of the NKG2D ligand RAET1G leads to cell surface expression of a glycosylphosphatidylinositol-linked isoform**. *J Biol Chem* 2010, **285**(22):16408-16415.

52. Hayakawa Y, Kelly JM, Westwood JA, Darcy PK, Diefenbach A, Raulet D, Smyth MJ: **Cutting edge: tumor rejection mediated by NKG2D receptor-ligand interaction is dependent upon perforin**. *J Immunol* 2002, **169**(10):5377-5381.

53. Yamaguchi K, Chikumi H, Shimizu A, Takata M, Kinoshita N, Hashimoto K, Nakamoto M, Matsunaga S, Kurai J, Miyake N *et al*: **Diagnostic and prognostic impact of serum-soluble UL16-binding protein 2 in lung cancer patients**. *Cancer Sci* 2012, **103**(8):1405-1413.

54. Nuckel H, Switala M, Sellmann L, Horn PA, Durig J, Duhrsen U, Kuppers R, Grosse-Wilde H, Rebmann V: **The prognostic significance of soluble NKG2D ligands in B-cell chronic lymphocytic leukemia**. *Leukemia* 2010, **24**(6):1152-1159.

55. Rothe A, Jachimowicz RD, Borchmann S, Madlener M, Kessler J, Reiners KS, Sauer M, Hansen HP, Ullrich RT, Chatterjee S *et al*: **The bispecific immunoligand ULBP2-aCEA redirects natural killer cells to tumor cells and reveals potent anti-tumor activity against colon carcinoma**. *Int J Cancer* 2014, **134**(12):2829-2840.

56. Li K, Mandai M, Hamanishi J, Matsumura N, Suzuki A, Yagi H, Yamaguchi K, Baba T, Fujii S, Konishi I: **Clinical significance of the NKG2D ligands, MICA/B and ULBP2 in ovarian cancer:**

high expression of ULBP2 is an indicator of poor prognosis. *Cancer Immunol Immunother* 2009, **58**(5):641-652.

57.   Raffaghello L, Prigione I, Airoldi I, Camoriano M, Levreri I, Gambini C, Pende D, Steinle A, Ferrone S, Pistoia V: **Downregulation and/or release of NKG2D ligands as immune evasion strategy of human neuroblastoma**. *Neoplasia* 2004, **6**(5):558-568.
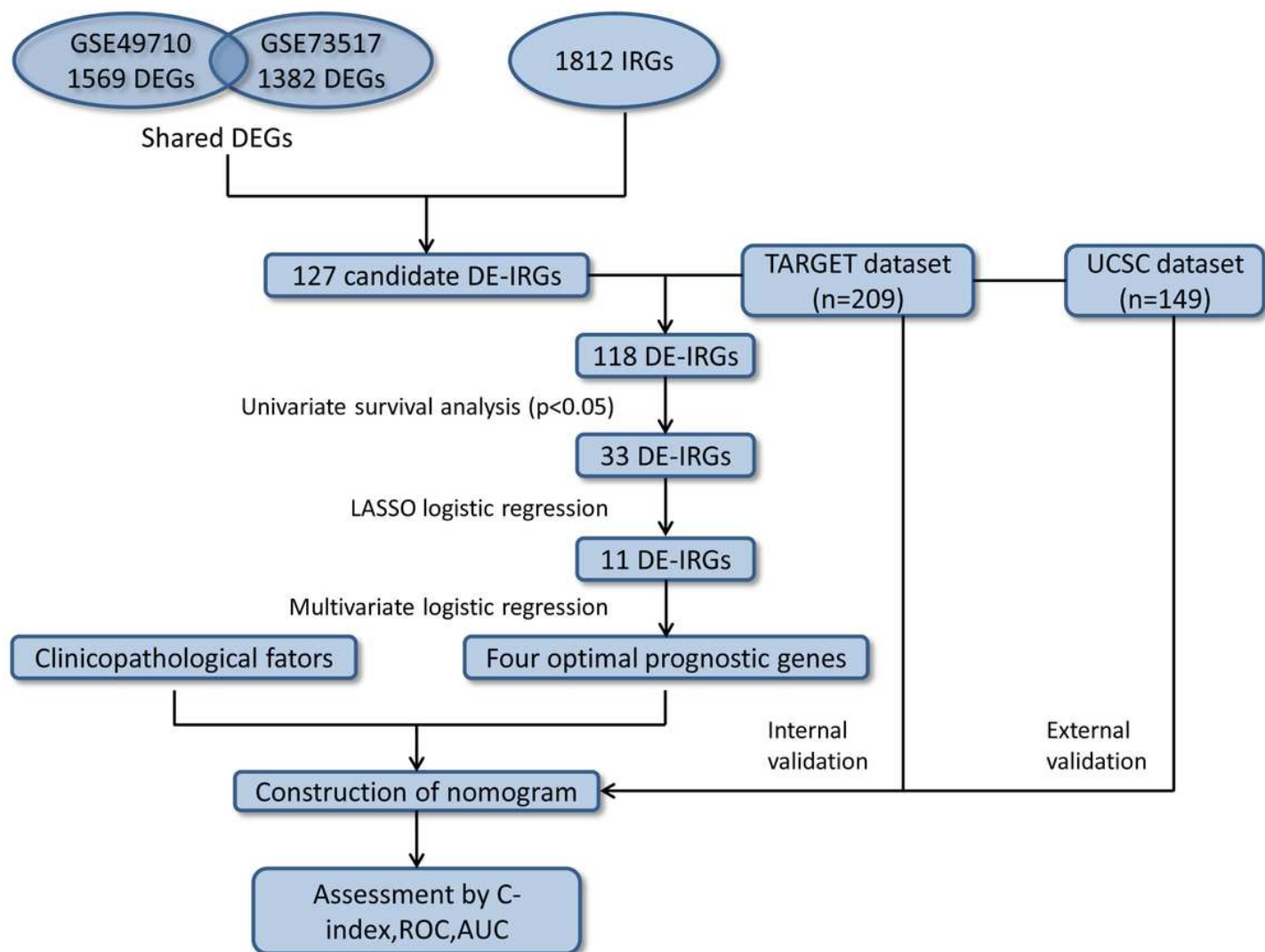
Figures

Figure 1

Overall design of the present study. DEGs differentially expressed genes, IRGs immune-related genes, DE-IRGs differentially expressed immune-related genes, TARGET the therapeutically applicable research to generate effective treatments, GEO gene expression omnibus, LASSO least absolute shrinkage and selection operator, C-index concordance index, ROC receiver operating characteristic, AUC area under the curve.
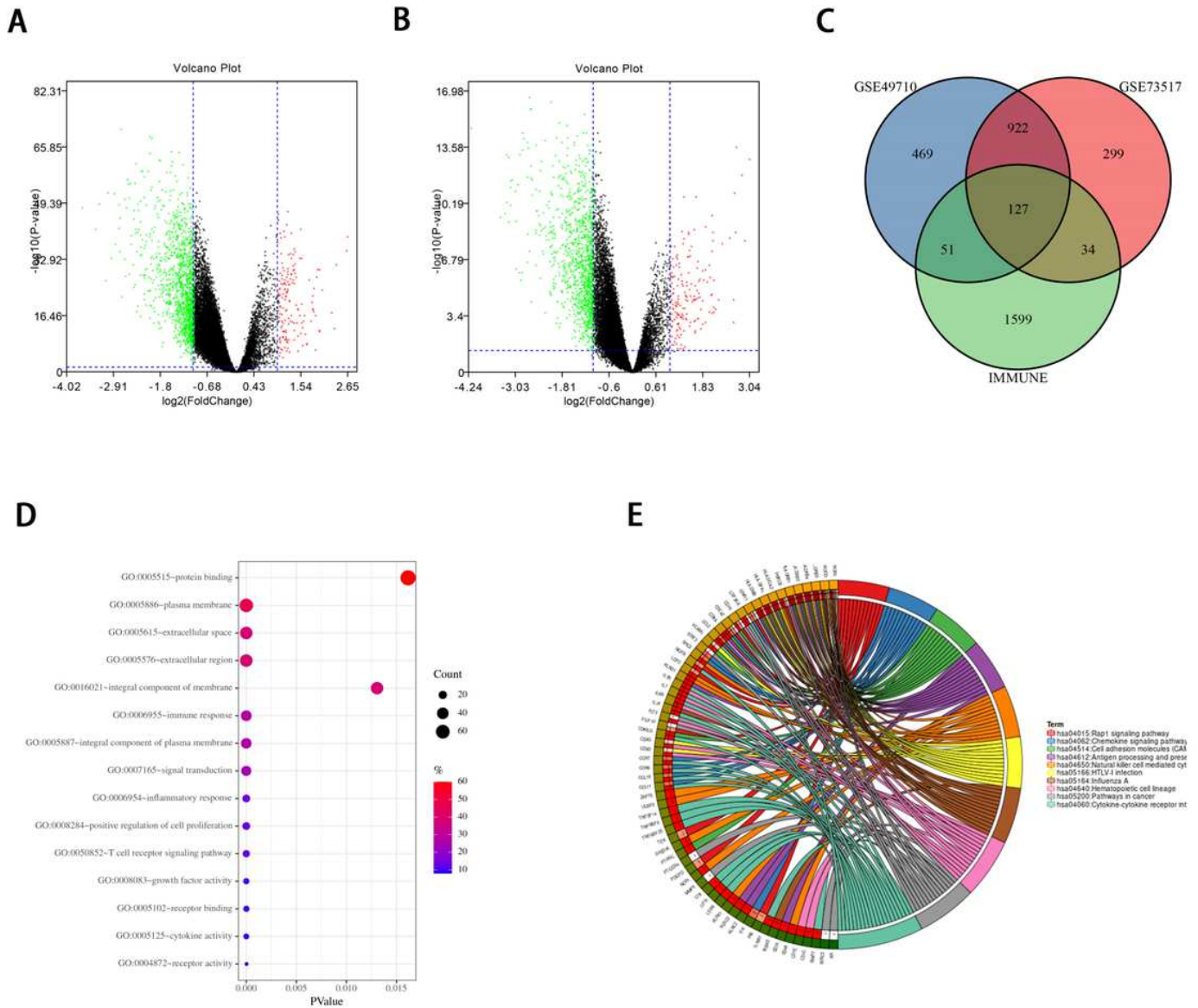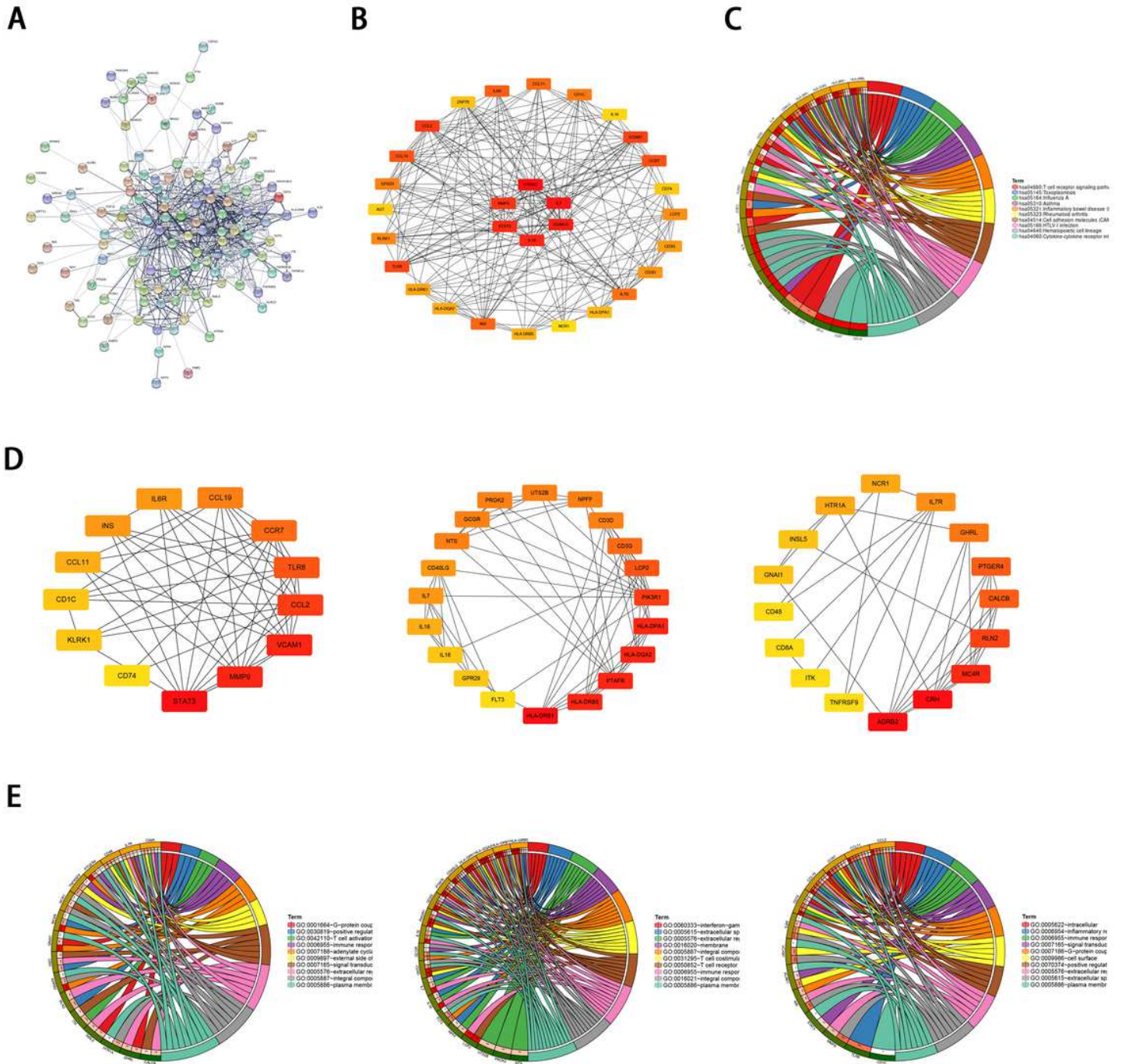
Figure 2

Identification of DE-IRGs and functional enrichment analysis. (A,B) Volcano plot illustrating differentially expressed genes (DEGs) between melanoma tissue and normal skin in GSE49710 (A) and GSE73517 (B). (C) Venn diagram of the overlapped genes between DEGs and IRGs. (D) Enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of DE-IRGs. (E) Enriched Gene Ontology (GO) pathways of DE-IRGs. MF molecular function, CC cell component, BP biological process.
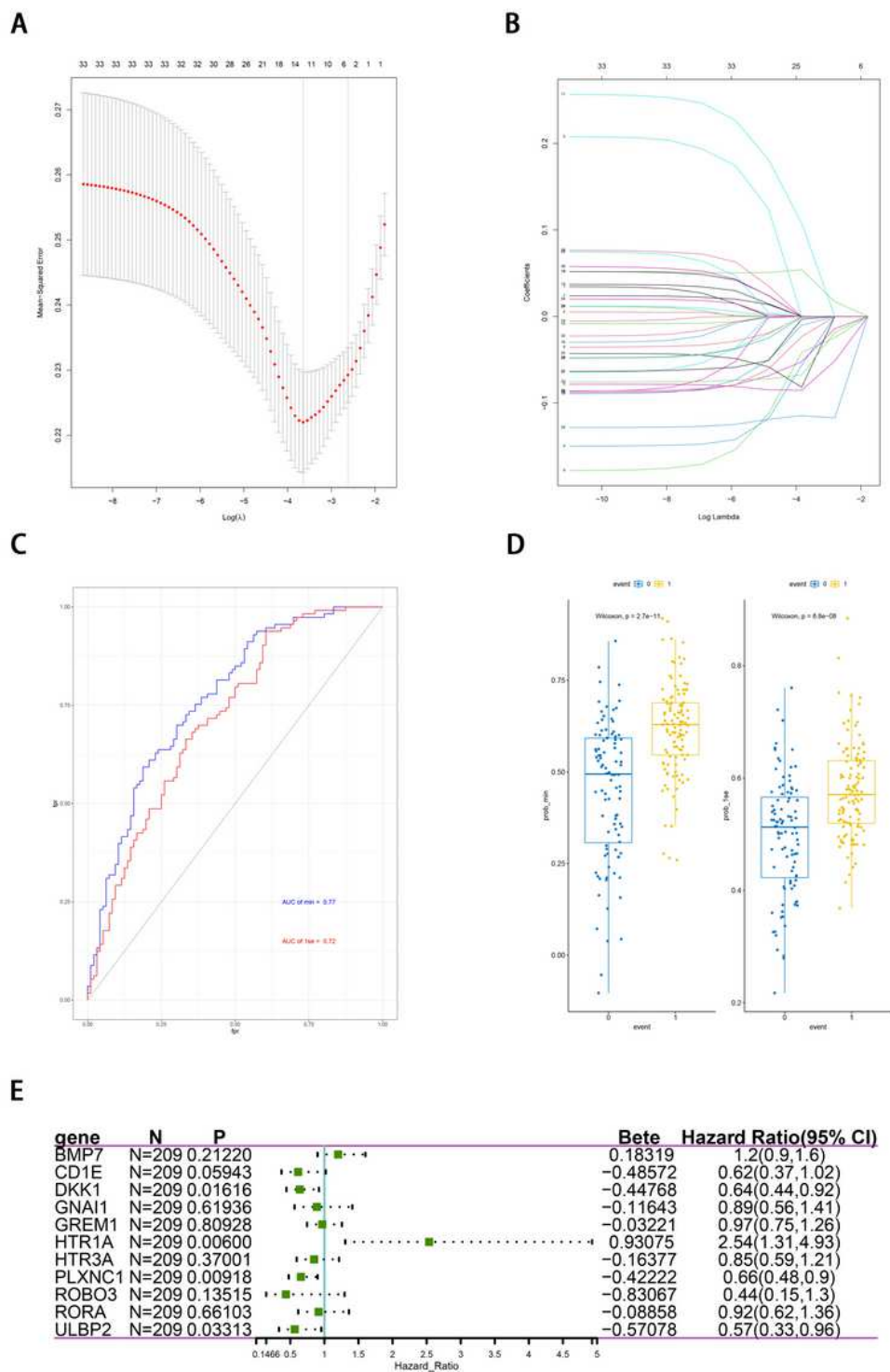
Figure 3

Functional analysis of IRG protein interaction. (A) the network analysisprotein-protein interaction PPI of IRG .(B) The top 30 gene network calculated based on the "degree" algorithm of the cytoscape hub plug-in. (C) Functional enrichment of top30 hub genes .(D) Three gene clusters based on the clustering algorithm of Cytoscape's MCODE plug-in. (E) Functional enrichment of MCODE gene cluster.

Figure 4

Identification and verification of prognostic gene signature. (A)(B) LASSO analysis for selecting the candidate IRGs in TCGA dataset. (C)(D) Comparison of the prediction performance of the two models ((λ min=0.02614 and λ 1se=0.07273). (E) Multivariate Cox regression analysis of 4 IRG.
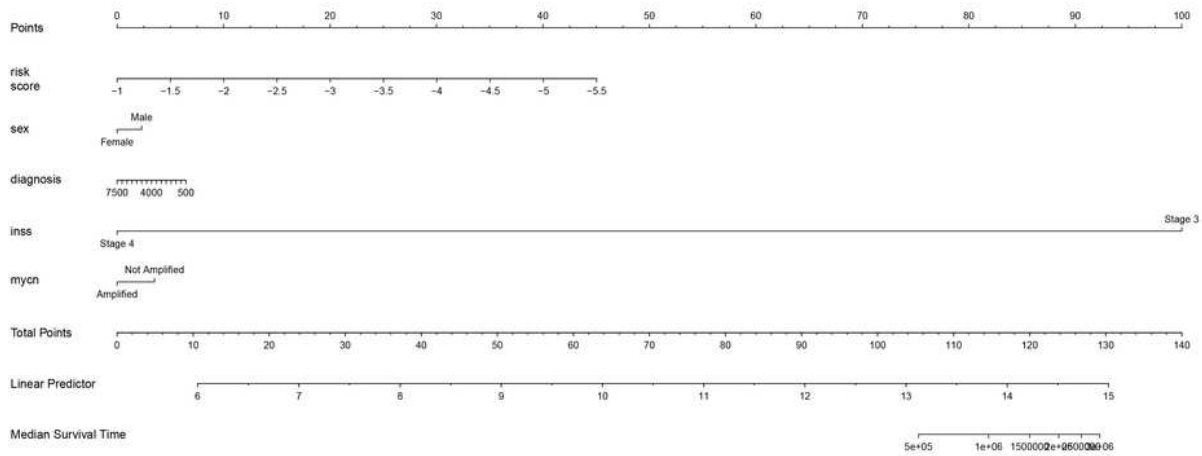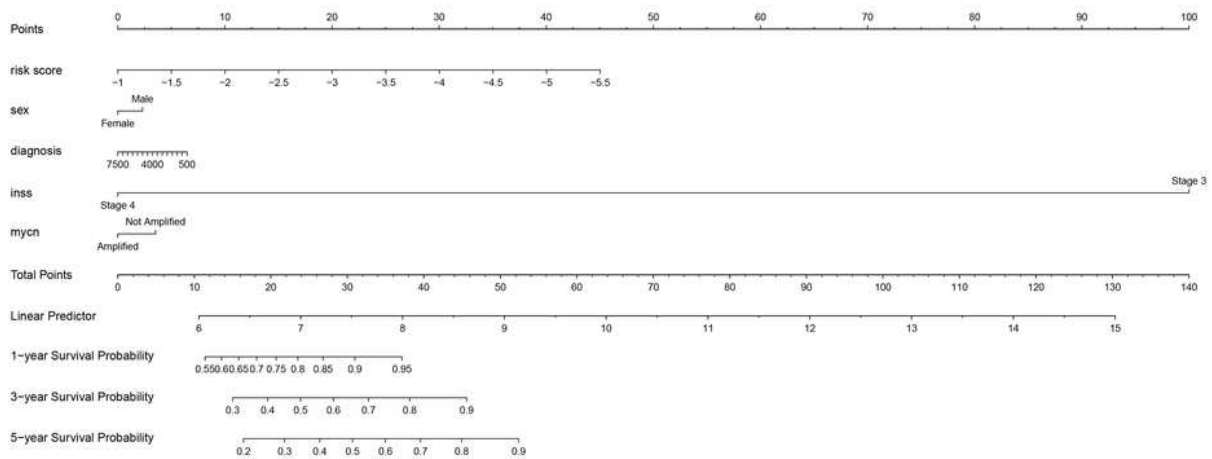
Figure 5

Prognostic analysis of the IRGs signature. (A)(B) (A,B) Kaplan−Meier survival curves of patients in high versus low risk groups in TARGET (A) and UCSC datasets (B). (C)(D) RCO curve of risk score in TARGET (C) and UCSC datasets (D). (E)(F) Distribution of IRGs score and survival days of each patient, and a heatmap of selected IRG expression profiles presented in order of IRGs score in TARGET (E) and UCSC datasets (F).
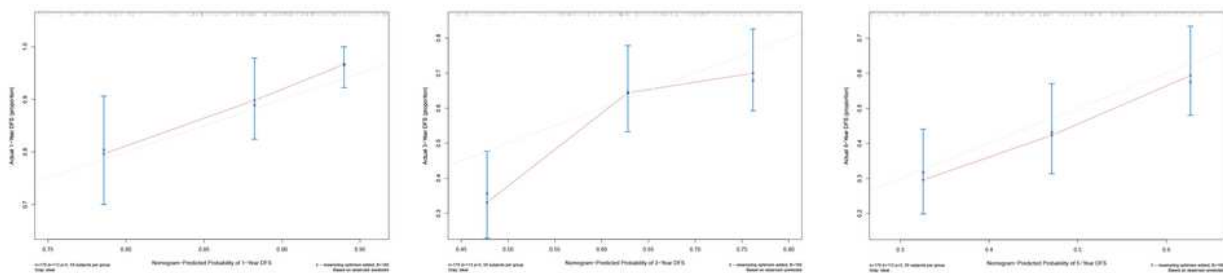
Figure 6

Draw and evaluate the nomogram combined with clinical pathological parameters. (A)(B) A nomogram of 1-year, 3-year, 5-year overall survival rate and median overall survival time. (C) The 1-, 3- and 5- calibration curves of the prediction model. The grey line represents a perfect prediction, and the red line describes the predictive performance of the nomogram, where the fitness of the black line to the grey line indicates a good prediction by the model.
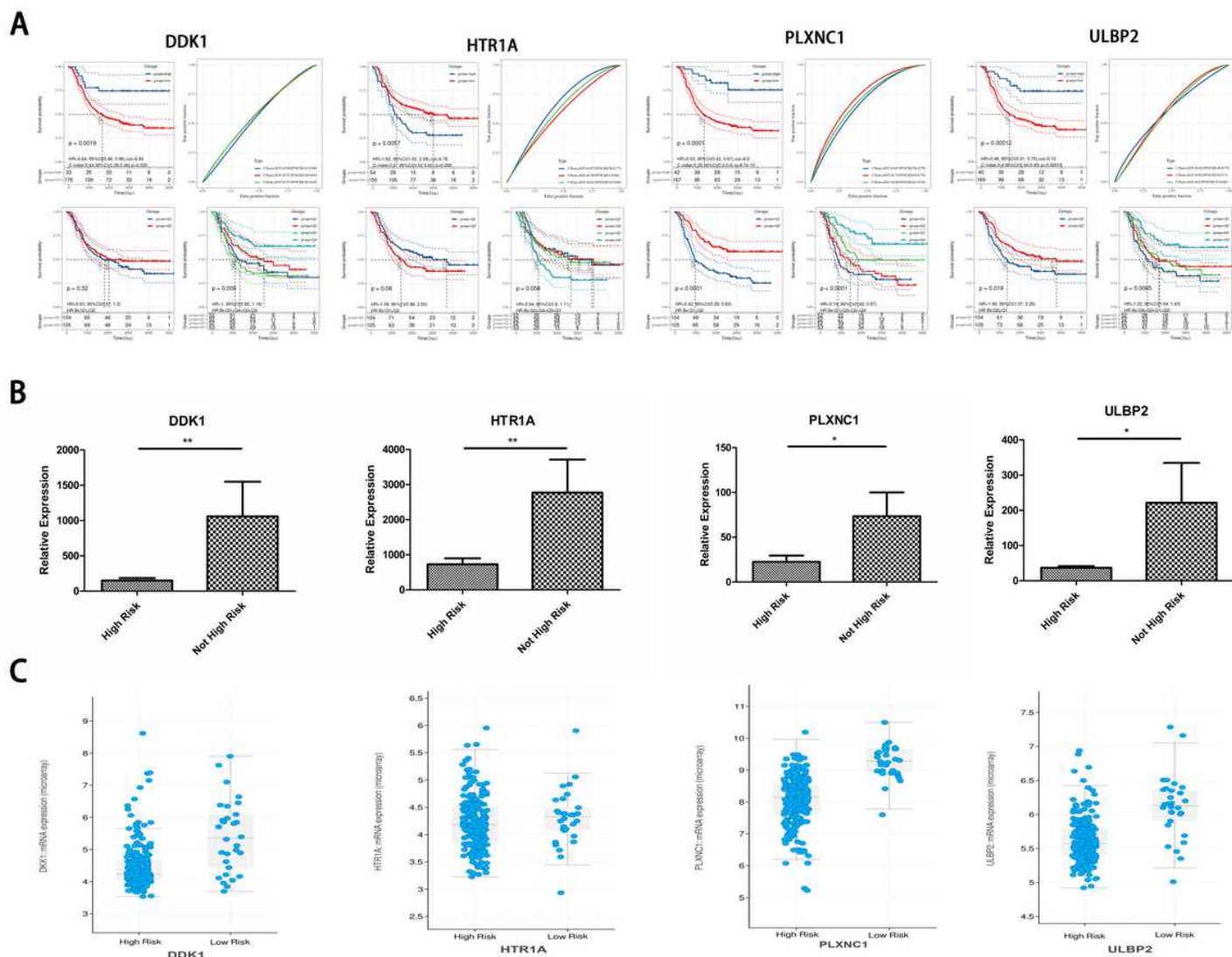
Figure 7

Verify the expression of DKK1, HTR1A, PLXNC1 and ULBP2 and evaluate their independent prediction performance. (A) Kaplan–Meier survival analysis and ROC curve of DKK1, HTR1A, PLXNC1 and ULBP2. (B) qRT-PCR detects the expression of DKK1, HTR1A, PLXNC1 and ULBP2 in the tissues of high-risk and non-high-risk patients. (C) The Cbioportal web server retrieves the expression differences of DKK1, HTR1A, PLXNC1 and ULBP2 in high-risk patients and high-risk patients.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigure1.tif
- SupplementaryFigure2.tif
- SupplementaryTableS1.pdf

- SupplementaryTableS2.pdf
- SupplementaryTableS3.pdf
- SupplementaryTableS4.pdf